

ANALISIS EMOSI PADA MEDIA SOSIAL TWITTER MENGUNAKAN METODE MULTINOMIAL NAÏVE BAYES DAN SYNTHETIC MINORITY OVERSAMPLING TECHNIQUE

Fritson Agung Julians Ayomi¹, Kania Evita Dewi²

^{1,2}Program Studi Teknik Informatika – Universitas Komputer Indonesia
Jl. Dipatiukur 112-116, Bandung
E-mail: kania.evita.dewi@email.unikom.ac.id ²

Abstrak

Media sosial Twitter sering digunakan untuk mengekspresikan emosi seseorang melalui sebuah cuitan. Penelitian tentang analisis emosi dalam media social twitter sudah banyak dilakukan. Mesin learning adalah tools yang banyak digunakan untuk melakukan pengkategorian emosi. Namun, ketidakseimbangan jumlah data antar kelas sering jadi masalah. Maka, penelitian ini bertujuan untuk mengetahui performansi hasil gabungan metode *Multinomial Naïve Bayes* (MNB) dan *Synthetic Minority Oversampling Technique* (SMOTE) untuk analisis emosi cuitan dari media sosial Twitter. Setiap cuitan melauai *preprocessing* data pada penelitian ini meliputi *case folding*, *data cleaning*, *convert slangword*, *convert negation*, *tokenization*, *stopword removal*, dan *stemming*. Untuk ekstraksi fitur digunakan metode n-gram dan untuk pembobotan fitur digunakan metode *term frequency*. Pengujian dilakukan menggunakan *K-Fold Cross Validation*. Berdasarkan hasil pengujian, dengan menggunakan SMOTE diperoleh rata-rata akurasi sebesar 0.65 atau 65% dan nilai rata-rata *f1-score* sebesar 0.66 atau 66%. Sedangkan tanpa SMOTE diperoleh rata-rata akurasi 0.64 atau 64% dan rata-rata *f1-score* sebesar 0.65 atau 65%. Walaupun dalam penelitian ini dapat ditunjukkan hasil dengan menggunakan SMOTE lebih baik 1% dalam pengkategorian emosi. Tetapi hasil yang diperoleh belum maksimal, masih perlu dikaji lagi untuk metode penyeimbangan data dan *mesin learning* yang lain.

Kata kunci : Analisis Emosi, *Synthetic Minority Oversampling Technique*, *Multinomial Naïve Bayes*

Abstract

Twitter social media is often used to express one's emotions through tweets. Much research has been conducted on emotional analysis in the social media Twitter. Machine learning is a tool that is widely used to categorize emotions. However, an imbalance in the amount of data between classes is often a problem. So, this research aims to determine the performance of the combined Multinomial Naïve Bayes (MNB) and Synthetic Minority Oversampling Technique (SMOTE) methods for emotional analysis of tweets from the social media Twitter. Each tweet through data preprocessing in this research includes case folding, data cleaning, convert slangword, convert negation, tokenization, stopword removal, and stemming. For feature extraction the n-gram method is used and for feature weighting the term frequency method is used. Testing was carried out using K-Fold Cross Validation. Based on the test results, using SMOTE an average accuracy of 0.65 or 65% was obtained and an average f1-score value of 0.66 or 66%. Meanwhile, without SMOTE, an average accuracy of 0.64 or 64% was obtained and an average f1-score of 0.65 or 65%. Although in this study it can be shown that the results using SMOTE are 1% better in categorizing emotions. However, the results obtained are not optimal, and other methods of data balancing and machine learning still need to be studied.

Keywords : *Emotion Analysis, Synthetic Minority Oversampling Technique, Multinomial Naïve Bayes*

1. PENDAHULUAN

Media sosial merupakan suatu platform digital untuk melakukan aktivitas sosial dalam jaringan internet. Saat ini banyak sekali media sosial yang digunakan oleh masyarakat Indonesia salah satunya adalah Twitter. Berdasarkan data dari <https://www.statista.com> [1], per Januari 2022 pengguna Twitter di Indonesia mencapai angka 18,45 juta. Dengan menggunakan media sosial Twitter, seseorang dapat mengekspresikan opini mereka melalui sebuah cuitan. Cuitan yang dibagikan pada media sosial twitter dapat mengandung emosi atau perasaan dari sang penulis. Emosi yang di deteksi dari setiap cuitan tersebut dapat digunakan

untuk pengambilan keputusan pada bidang bisnis, hukum, kesehatan atau dalam kehidupan bersosial. Seperti pada penelitian Hilda Rachmi, Suparni, dan Ahmad Al Kaafi [2], dilakukan analisis emosi mengenai opini masyarakat terhadap diterapkannya sistem ganjil genap di kota Bogor melalui postingan masyarakat di sosial media.

Penelitian mengenai analisis emosi saat ini telah banyak dilakukan. Salah satunya adalah penelitian yang dilakukan oleh Fildza Sakinah Alnaz dan Warih Maharani [3]. Pada penelitian tersebut dilakukan klasifikasi emosi pada cuitan pengguna di media sosial Twitter dengan membandingkan peformansi antara kombinasi metode *Naïve Bayes* dengan fitur N-Gram dan *Naïve Bayes* dengan fitur TF-IDF. Dimana akurasi terbaik yang didapatkan sebesar 0.6485 dengan rata-rata nilai *precision* 0.69, *recall* 0.64, dan *f1-score* 0.66. Dalam penelitian Fildza menggunakan lima emosi yaitu *anger*, *happy*, *sadness*, *fear*, dan *love*. Terdapat masalah pada penelitian tersebut yaitu adanya ketidakseimbangan data antar kelas, dimana jumlah emosi *anger*, *happy*, *sadness*, *fear*, dan *love* secara berturut-turut yaitu 1101, 1017, 997, 649, dan 637. Data yang sama digunakan dalam penelitian ini. Masalah ketidakseimbangan data cukup sering terjadi dalam masalah klasifikasi karena pada dasarnya data yang diambil langsung dari sumber data merupakan data yang tidak seimbang sehingga membuat proses klasifikasi menjadi lebih menantang [4]. Untuk mengatasi ketidakseimbangan data dapat dilakukan proses *oversampling*. Pada penelitian Afifah Nurlaila, Wiranto, dan Ristu Saptono [5] mengenai klasifikasi emosi dari pelanggan produk perawatan kulit dari cuitan Twitter menggunakan metode *Naïve Bayes*, dilakukan proses *oversampling* menggunakan *Random Over Sampling* (ROS) karena saat pengujian terdapat satu kelas yang memiliki jumlah data yang lebih banyak dari kelas lainnya sehingga menghasilkan nilai *precision* dan *recall* yang sangat rendah mencapai 0%. Proses *oversampling* yang dilakukan pada penelitian tersebut dapat meningkatkan nilai *recall*, *precision*, dan *f1-measure*. Namun, dikarenakan cara kerja metode ROS yang menduplikasi data yang sama persis dari data kelas minoritas sehingga mungkin menyebabkan *overfitting* [6].

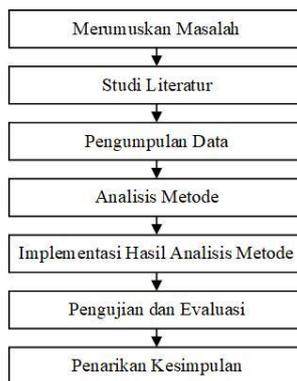
Synthetic Minority Oversampling Technique (SMOTE) merupakan salah satu perkembangan dari metode *oversampling*. Teknik SMOTE mirip dengan ROS, namun sampel baru yang dihasilkan oleh SMOTE dibuat dengan menerapkan konsep *Nearest Neighbour* [6]. Saat ini telah banyak penelitian yang menerapkan metode SMOTE. Pada penelitian Cristian Padurariu dan Mihaela Elena Breaban [7] dalam melakukan perbandingan performansi pada kombinasi antara berbagai metode pembelajaran mesin, *sampling*, dan pembobotan kata untuk menghadapi ketidakseimbangan data didapatkan bahwa metode *oversampling* SMOTE mampu memberikan performansi yang sama bahkan melebihi metode ROS pada beberapa kasus pengujian. Penelitian yang dilakukan oleh Saifurrachman Chohan, Arifin Nugroho, Achmad Maezar Bayu Aji, dan Windu Gata [8] dalam melakukan analisis sentimen pada aplikasi Duolingo menggunakan metode *Naïve Bayes* dan SMOTE mampu memberikan performansi yang sangat baik dimana akurasi yang dihasilkan sebesar 91.95% dan 0.740 untuk AUC.

Berdasarkan uraian latar belakang diatas maka pada penelitian ini akan dilakukan penerapan metode SMOTE untuk mengatasi ketidakseimbangan data pada analisis emosi cuitan dari media sosial Twitter dan menggunakan *Naïve Bayes* dengan model *Multinomial Naïve Bayes* sebagai metode klasifikasi yang akan digunakan.

2. METODOLOGI

2.1 Metode Penelitian

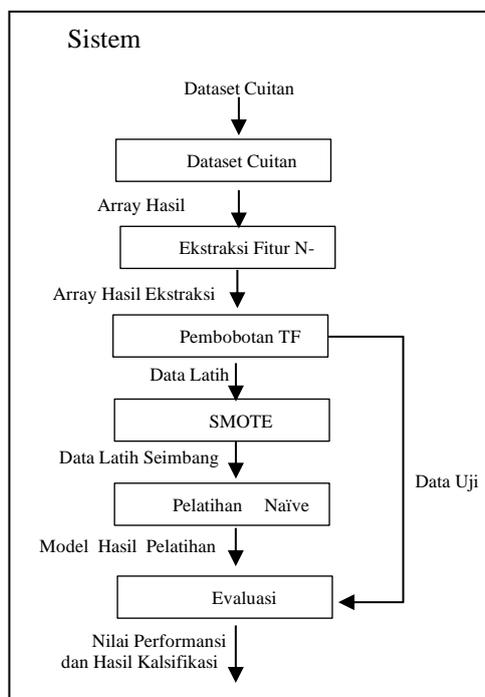
Metode penelitian yang digunakan pada penelitian ini adalah metode penelitian kuantitatif. Tahapan awal dalam penelitian ini dirumuskan masalah yang dibahas dalam penelitian ini. Untuk menyelesaikan masalah yang sudah dirumuskan peneliti melakukan studi Pustaka untuk mencari solusi yang terbaik untuk menyelesaikan masalah. Untuk menguji solusi hasil studi literatur peneliti mengumpulkan data yang akan diproses. Kemudian mencoba menganalisis data dan metode yang digunakan. Hasil dari analisis, diimplementasikan dengan membangun sebuah simulator. Untuk membuktikan teori dan implementasi berjalan sesuai dilakukan pengujian dan evaluasi. Penarikan kesimpulan adalah tahap akhir dalam penelitian ini. Adapun tahapan dari metode penelitian yang digunakan dapat dilihat pada gambar 1.



Gambar 1. Tahapan Metode Penelitian

2.2 Gambaran Umum Sistem

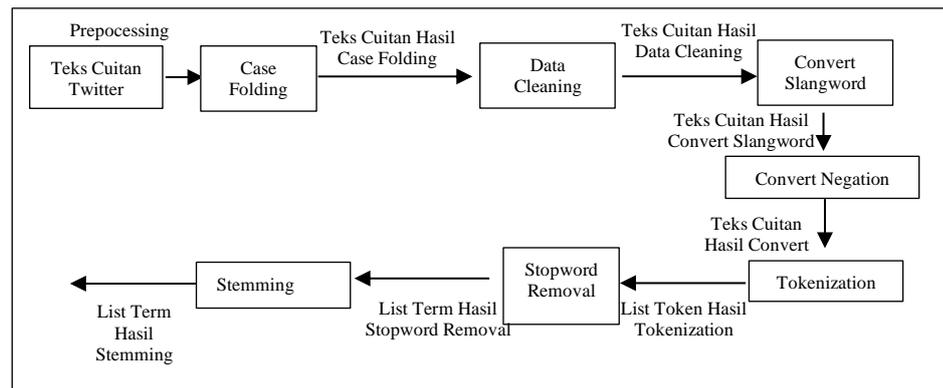
Sistem analisis emosi dari cuitan Twitter pada penelitian ini terbagi menjadi 7 tahap yaitu *preprocessing*, *split data*, ekstraksi fitur n-gram, pembobotan *term frequency*, SMOTE, pelatihan *Naïve Bayes*, dan evaluasi. Urutan dari tahapan pada sistem dapat dilihat dari gambaran umum sistem pada gambar 2 berikut.



Gambar 2. Gambaran Umum Sistem

2.2.1 Preprocessing

Dalam penelitian ini ada 7 tahapan metode *preprocessing* yaitu *case folding*, *data cleaning*, *convert slangword*, *convert negation*, *tokenization*, *stopword removal*, dan *stemming*. Urutan dari tahapan *preprocessing* dapat dilihat pada gambar 3 berikut.



Gambar 3. Tahapan Preprocessing

1. Case Folding
Case folding merupakan proses mengubah semua huruf kapital pada sebuah dokumen teks menjadi huruf kecil [11]. Proses ini dilakukan karena data teks yang dimiliki tidak selalu terstruktur dan konsisten dalam penggunaan huruf kapital.
2. Data Cleaning
Data cleaning merupakan proses menghilangkan atau mengubah hal-hal yang dianggap sebagai *noise* dari sebuah data teks [12] dan hanya menyisakan karakter berupa huruf alfabet. *Noise* tersebut dapat berupa karakter seperti tanda baca, angka, simbol atau karakter alfabet yang terdiri dari satu hingga beberapa karakter dan dianggap hanya mengganggu suatu data teks.
3. Convert Slangword
Slangword atau "kata gaul" merupakan istilah-istilah atau bahasa yang sering digunakan dalam berkomunikasi pada saat ini [3]. *Convert slangword* dilakukan untuk mengubah *slang* yang dapat berupa singkatan atau kata yang tidak baku menjadi kata utuh atau kata baku. Kamus kata yang digunakan pada penelitian ini berasal dari penelitian Mei Silviana Saputri, Rahmad Mahendra, dan Mirna Adriani [9].
4. Convert Negation
Convert negation merupakan proses penggabungan kata negasi dengan kata selanjutnya. Kata negasi seperti kata "bukan", "jangan", "tidak" dan sebagainya dapat mengubah makna sentimen dari sebuah dokumen [13].
5. Tokenization
Tokenization merupakan sebuah proses pemecahan kalimat menjadi karakter atau potongan kata, dimana hasil dari pemecahan disebut dengan istilah token [14].
6. Stopword Removal
Stopword Removal merupakan proses menghilangkan kata-kata yang tidak relevan, kata yang tidak memiliki makna tersendiri atau kata yang sering muncul dan menjadi tidak penting [15], seperti kata "dan", "pada", "pula", "saat" serta kata-kata lain yang terdapat dalam *stopword list*. *Stopword list* pada penelitian ini didapatkan dari *library Natural Language Toolkit (NLTK)*.
7. Stemming
Stemming merupakan proses menghilangkan imbuhan yang terdiri dari awalan, akhiran, awalan dan akhiran, dan sisipan dari sebuah kata untuk mendapatkan kata dasarnya [15]. Proses *stemming* pada penelitian ini dilakukan menggunakan *library Sastrawi*.

2.2.2 Splitting data

Dataset yang digunakan merupakan cuitan berbahasa Indonesia dari media sosial Twitter dengan format .csv yang diambil dari penelitian Mei Silviana Saputri, Rahmad Mahendra, dan Mirna Adriani [9]. Total cuitan pada dataset tersebut berjumlah 4401 cuitan yang terbagi menjadi 5 kelas emosi yaitu *anger*, *happy*, *sadness*, *fear*, dan *love*. Jumlah data dari setiap kelas emosi dapat dilihat pada tabel 1 berikut.

Tabel 1. Detail Dataset

Label Emosi	Jumlah Data Cuitan
Anger	1101
Happy	1017
Sadness	997
Fear	649
Love	637

Dari tabel 1 diketahui bahwa ada ketidakseimbangan jumlah data antar kelas emosi, ketidakseimbangan data tersebut dapat mempengaruhi hasil dari proses klasifikasi. Dimana kelas emosi dengan jumlah data paling banyak (mayoritas) akan memberikan hasil klasifikasi yang lebih baik dari pada kelas emosi dengan jumlah data yang sedikit (minoritas) [10]. Contoh sampel dari cuitan media sosial Twitter dari dataset yang digunakan dapat dilihat pada tabel 2 berikut.

Tabel 2. Contoh Cuitan Dari Dataset

No	Label	Cuitan
1	Anger	Bikin status d sosmed cm ngejelekin pacarna-cewek macam apaan itu.....giliran d beri hadiah di ajak jalan kga d bikin status..... *Muntahin ajan
2	Fear	Nah.... Yg saya takutkan itu mas. Nanti di tumpangi politik dibalut agama. Itu yg bahaya, soale masyarakat kita sedang sensitif dgn hal tersebut.
3	Happy	Antri di bank situasyen: mbak dapet nomor antrian berapa ? 227 buk , ini ambil aja nomor antrian biar cepet ---200 *dalamhati* rejeki anak baik
4	Love	Setiap kesempatan yg pernah hadir tuk dapat membuatmu selalu merasa aman dan nyaman, kini jadi suatu kehormatan yg pernah didapat
5	Sadness	Nggada sinyal di rumah, membuatku banyak berfikir, berfikir doank tapi, pikiran ke mana mana, random.. Tapi pasti ada dia :(

Dilihat dari contoh sampel data cuitan Twitter pada tabel 2 diketahui bahwa dataset yang digunakan memiliki banyak *noise* yang dapat mempengaruhi performansi klasifikasi. Hal ini dapat terlihat dimana terdapat ketidaksesuaian penggunaan huruf besar dan kecil, ketidaksesuaian penggunaan tanda baca, *slang* yang berupa singkatan atau istilah, dan kata yang tidak memiliki makna penting. Berdasarkan ini lah *preprocessing* pada tahapan sebelumnya dilakukan. Pada tahapan kedua ini dilakukan *spliting data* diperlukan mengingat *machine learning* yang digunakan bersifat *supervised*.

2.2.3 N-Gram

Metode n-gram digunakan untuk mendapatkan potongan karakter berupa huruf atau kata sejumlah *n* dari teks sumber secara terus-menerus hingga akhir dokumen [16]. N-gram dapat dibedakan berdasarkan dari jumlah potongan karakter sebesar *n*. Misalnya jika *n* = 1 maka disebut dengan *unigram*, *n* = 2 disebut bigram, *n* = 3 disebut dengan *trigram* dan seterusnya. Pada penelitian ini kombinasi dari *unigram*, *bigram*, dan *trigram* akan digunakan sebagai fitur.

2.2.4 Term Frequency

Term Frequency merupakan suatu metode pembobotan yang menentukan bobot dari suatu term di dalam suatu dokumen. Didalam penelitian ini digunakan metode *Raw Term Frequency*, dimana pembobotan diperoleh dari jumlah kemunculan *term* dalam suatu dokumen [17]. Maksudnya jika kata “status” didalam suatu cuitan muncul lima kali, maka *term frequency* untuk kata “status” pada cuitan tersebut bernilai lima.

2.2.5 Synthetic Minority Over-sampling Technique

Synthetic Minority Over-sampling Technique (SMOTE) merupakan salah satu perkembangan dari metode *oversampling*. SMOTE menambah jumlah data kelas minoritas dengan cara mengambil sampel dari kelas minoritas dan membangkitkan data buatan berdasarkan *k-nearest neighbour* dari kelas minoritas [18]. Metode SMOTE bekerja dengan melakukan pengelompokan data berdasarkan jarak *ecludian* antar data. Jumlah replikasi data pada kelas minoritas disesuaikan berdasarkan jumlah data pada kelas mayoritas, dimana jumlah replikasi data harus sama dengan jumlah nilai *k* pada *nearest neighbour*, jika jumlah replikasi data yang dilakukan sebanyak *n* maka jumlah *k* harus sebanyak *n-1* [19]. Algoritma dari metode SMOTE dapat dilihat melalui *pseudocode* [18] pada tabel 3 berikut.

Tabel 3. Pseudocode SMOTE

Algoritma SMOTE(T, N, k)
<p>Masukan: Jumlah sampel kelas minoritas T; Jumlah SMOTE $N\%$; Jumlah tetangga terdekat k</p> <p>Keluaran: $(N/100) * T$ sampel kelas minoritas sintetis (*Jika N kurang dari 100%, acak sampel kelas minoritas karena hanya persentase acak dari data yang akan di SMOTE.*)</p> <p>If $N < 100$ then Acak sampel kelas minoritas T $T = (N/100) * T$ $N = 100$ endif</p> <p>$N = (\text{int})(N/100)$ (*Jumlah SMOTE diasumsikan dalam kelipatan integral dari 100.*)</p> <p>$k =$ Jumlah tetangga terdekat $\text{numattrs} =$ Jumlah atribut $\text{Sample}[\][\]:$ array untuk sampel kelas minoritas asli $\text{newindex}:$ menghitung jumlah sampel sintetis yang dihasilkan, diinisialisasi ke 0 $\text{Synthetic}[\][\]:$ array untuk sampel sintetis (*Hitung k tetangga terdekat hanya untuk setiap sampel kelas minoritas.*)</p> <p>for $i \leftarrow 1$ to T Hitung k tetangga terdekat untuk i, dan simpan indeks di nnarray $\text{Populasi}(N, i, \text{nnarray})$ endfor</p> <p>$\text{Populasi}(N, i, \text{nnarray})$ (*Fungsi untuk menghasilkan sampel sintetis.*)</p> <p>while $N \neq 0$ Pilih nomor acak antara 1 dan k, sebut saja nn. Langkah ini memilih salah satu k tetangga terdekat dari i. for $\text{attr} \leftarrow 1$ to numattrs Hitung: $\text{dif} = \text{Sample}[\text{nnarray}[\text{nn}]][\text{attr}] - \text{Sample}[i][\text{attr}]$ Hitung: $\text{gap} =$ angka acak antara 0 dan 1 $\text{Synthetic}[\text{newindex}][\text{attr}] = \text{Sample}[i][\text{attr}] + \text{gap} * \text{dif}$ endfor $\text{newindex}++$ $N = N - 1$ endwhile</p> <p>return (*Akhir Populasi.*)</p> <p>Akhir dari Pseudo-Code.</p>

2.2.6 Naïve Bayes

Naïve Bayes merupakan metode klasifikasi yang berasal dari penerapan teorema bayes. Asumsi yang sangat kuat (naif) terhadap independensi dari setiap kondisi atau kejadian merupakan ciri utama dari metode Naïve Bayes [16]. Persamaan dari perhitungan posterior dan prior dapat dilihat pada persamaan (1) dan (2) berikut [20].

$$P(A_i|B) = P(B|A_i) \times P(A_i) = \left(\prod_{k=1}^n P(b_k|A_i) \right) P(A_i) \quad (1)$$

Dimana:

$P(A_i|B)$ = Probabilitas dokumen B termasuk kelas A_i (posterior).

$P(A_i)$ = Probabilitas prior kelas A_i .

b_k = Term b_k pada dokumen.

$P(b_k|A_i)$ = Probabilitas term b_k dari kelas A_i .

$$P(A_i) = \frac{n(A_i)}{N} \quad (2)$$

Dimana:

- $P(A_i)$ = Probabilitas prior kelas A_i .
- $n(A_i)$ = Jumlah kelas A_i pada seluruh dokumen.
- N = Jumlah seluruh dokumen (banyaknya data latih).

Model *Naïve Bayes* yang digunakan pada penelitian ini adalah *Multinomial Naïve Bayes*. Dikarenakan data hasil proses pada penelitian ini merupakan data kontinu maka digunakan persamaan densitas *Gaussian* untuk menghitung nilai probabilitas dari kata atau *term*. Persamaan dari densitas *Gaussian* dapat dilihat pada persamaan (3) berikut [21].

$$P(b_k|A_i) = \frac{1}{\sqrt{2\pi\sigma_{ik}}} e^{-\frac{(b_k-\mu_{ik})^2}{2\sigma_{ik}^2}} \quad (3)$$

Dimana:

- b_k = Nilai *term* b_k .
- σ_{ik} = Standar deviasi *term* b_k pada kelas A_i .
- μ_{ik} = Rata-rata hitung (*mean*) *term* b_k pada kelas A_i .
- π = Phi 3.14

Persamaan untuk melakukan perhitungan rata-rata dan standar deviasi dapat dilihat pada persamaan (4) dan (5) berikut [21].

$$\mu_i = \frac{\sum_{i=1}^n x_i}{n} \quad (4)$$

Dimana:

- μ_i = Rata-rata hitung (*mean*) *term* ke- i .
- x_i = Nilai *term* ke- i .
- n = Jumlah *term* ke- i .

$$\sigma_i = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}} \quad (5)$$

Dimana:

- σ_i = Standar deviasi *term* ke- i .
- x_i = Nilai *term* ke- i .
- μ = Rata-rata hitung (*mean*) *term* ke- i .
- n = Jumlah *term* ke- i .

2.2.7 Confusion Matrix

Confusion Matrix merupakan metode yang digunakan untuk menganalisis peformansi dari metode klasifikasi dalam melakukan pengenalan kelas [15]. *Confusion matrix* memiliki beberapa variable yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Dengan menggunakan *confusion matrix* dapat dilakukan evaluasi dari metode klasifikasi terhadap nilai *accuracy*, *precision*, *recall*, dan *f1-Score*. Untuk melakukan perhitungan nilai *accuracy*, *precision*, *recall*, dan *f1-score* dapat digunakan rumus pada persamaan (6) hingga (9) [22].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

$$F1 - Score = 2 \times \frac{recall \times precision}{recall + precision} \quad (9)$$

2.2.8 K-fold Cross Validation

K-fold Cross Validation merupakan salah satu jenis dari pengujian *Cross Validation* yang digunakan untuk mengevaluasi performansi dari sebuah metode algoritma dengan cara membagi kumpulan data secara

acak dan mengelompokkan data tersebut sebanyak nilai k pada K-fold [23]. Pada metode *K-fold Cross Validation* dataset akan dibagi menjadi beberapa partisi secara acak, dimana data partisi tersebut diproses sebanyak k kali iterasi prosedur dengan setiap iterasi prosedur menggunakan data dari partisi ke- k sebagai data uji dan menggunakan data lain yang tersisa sebagai data latih [23]. Pada penelitian ini *K-fold cross validation* digunakan dalam proses klasifikasi dengan jumlah *fold* sebanyak 10-*fold*.

3. HASIL DAN PEMBAHASAN

Skenario pengujian dalam penelitian ini terdiri dari 5 skenario, yaitu: 1. Data latih dioversampling menggunakan SMOTE, kemudian diklasifikasi dengan metode Gaussian Naïve Bayes. 2. Data tidak dioversampling (menggunakan data asli), langsung diklasifikasi dengan metode Gaussian Naïve Bayes. 3. Data latih dioversampling, kemudian diklasifikasi menggunakan Multinomial Naïve Bayes. 4. Data latih tidak dioversampling (menggunakan data asli), setelah pembobotan diklasifikasi menggunakan Multinomial Naïve Bayes. 5. Karena didalam SMOTE menggunakan metode KNN, maka skenario kelima ini adalah untuk melihat k terbaik yang harus digunakan untuk proses SMOTE. Skenario ke-1 dan ke-3 dilakukan untuk mengetahui algoritma mana yang lebih baik dalam pengambilan kesimpulan. Sedangkan skenario ke-1 dan ke-4 untuk sebagai standar untuk melihat pengaruh SMOTE dan metode klasifikasi mana yang terbaik. Kelima skenario ini dapat diringkas dalam tabel 4

Tabel 4. Skenario Pengujian Performansi

No	Skenario Pengujian Performansi
1	<i>Gaussian Naïve Bayes-SMOTE</i>
2	<i>Gaussian Naïve Bayes</i>
3	<i>Multinomial Naïve Bayes-SMOTE</i>
4	<i>Multinomial Naïve Bayes</i>
5	Nilai k pada SMOTE ($k = 1, 3, 5, 7, 9$)

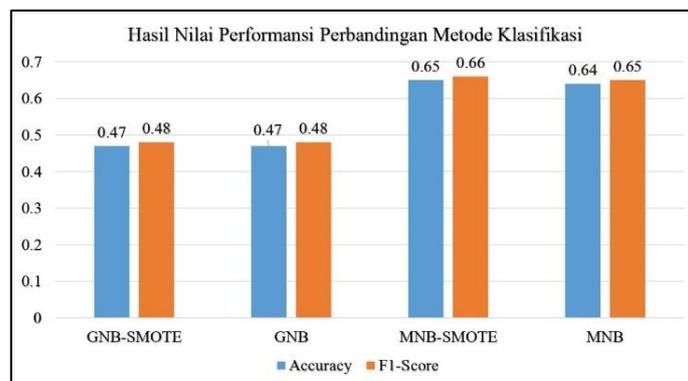
3.1 Hasil

Pertama-tama akan dijabarkan hasil dari proses *oversampling* data latih menggunakan metode SMOTE. Perbandingan dari jumlah data latih sebelum dan sesudah dilakukan proses *oversampling* SMOTE dapat dilihat pada tabel 5 berikut.

Tabel 5. Perbandingan Jumlah Data Latih Sebelum Dan Sesudah SMOTE

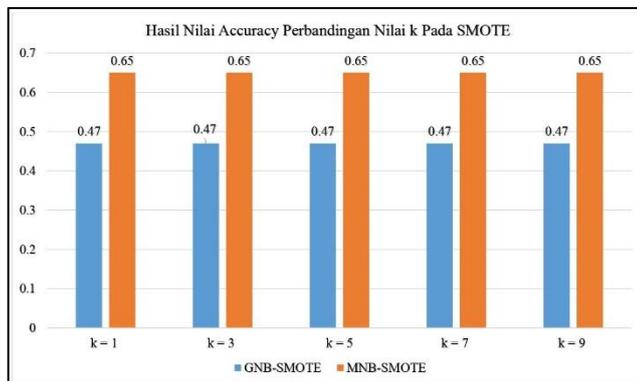
Kelas	Jumlah Data Latih Sebelum SMOTE	Jumlah Data Latih Sesudah SMOTE
Anger	990	990
Happy	916	990
Sadness	897	990
Fear	584	990
Love	573	990

Selanjutnya adalah hasil dari pengujian performansi klasifikasi emosi terhadap skenario pengujian nomor 1 sampai 4 yang telah dijabarkan pada tabel 4. Hasil pengujian tersebut dapat dilihat pada gambar 4 berikut.

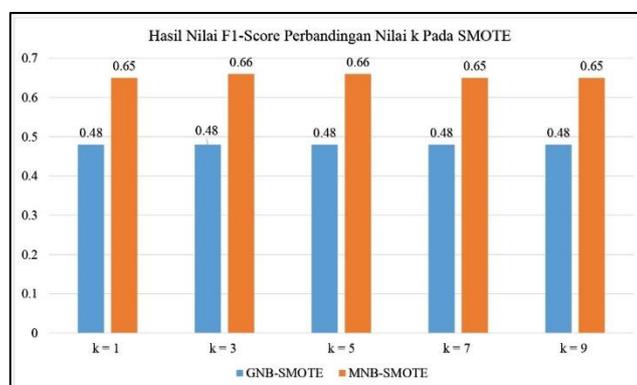


Gambar 4. Grafik Perbandingan Nilai Performansi Metode Klasifikasi

Terakhir adalah hasil dari pengujian pengaruh nilai k pada SMOTE terhadap performansi proses klasifikasi yang dapat dilihat pada gambar 5 dan 6 berikut.



Gambar 5. Grafik Perbandingan Nilai Akurasi Nilai k Pada SMOTE



Gambar 6. Grafik Perbandingan Nilai F1-Score Nilai k Pada SMOTE

Pada tabel 6 diperlihatkan hasil salah satu fold dari skenario pengujian 3 dan 4. Dapat dilihat pada tabel 6 hampir semua label mengalami kenaikan baik akurasi maupun f1 setelah menggunakan SMOTE kecuali “fear”. Tetapi jika secara TP, yaitu data yang terprediksi benar label “fear”, meningkat setelah di SMOTE.

Tabel 6. Hasil Klasifikasi menggunakan SMOTE dan tidak menggunakan SMOTE menggunakan Multinomial Naive-Bayes

Emosi	Menggunakan SMOTE						Tanpa SMOTE					
	TP	TN	FP	FN	Akurasi	F-1	TP	TN	FP	FN	Akurasi	F-1
Anger	86	288	42	25	0.85	0.72	89	279	51	22	0.83	0.71
Fear	40	366	10	25	0.92	0.70	36	374	2	29	0.93	0.70
Happy	66	305	35	35	0.84	0.65	66	301	39	35	0.83	0.64
Love	46	365	12	18	0.93	0.75	37	369	8	27	0.92	0.68
Sadness	58	295	46	42	0.80	0.57	60	288	53	40	0.79	0.56

3.2 Pembahasan

Berdasarkan hasil pengujian yang telah dilakukan diketahui bahwa metode SMOTE dapat bekerja dengan baik dalam melakukan *oversampling* data latih yang tidakseimbang, dimana pada penelitian ini SMOTE mampu membuat jumlah data latih dari kelas minoritas *love*, *fear*, *sadness*, dan *happy* menjadi sebanyak jumlah data latih dari kelas mayoritas *anger*. Dari grafik hasil pengujian performansi pada gambar 4 dapat dikatakan bahwa performansi dari gabungan metode *Naive Bayes* dan SMOTE dalam melakukan analisis emosi pada cuitan Twitter masih kurang baik. Dimana nilai performansi tertinggi didapatkan pada metode gabungan *Multinomial Naive Bayes* (MNB) dan SMOTE dengan nilai akurasi rata-rata sebesar 0.65 atau 65% dan nilai *f1-score* rata-rata sebesar 0.66 atau 66%. Dan dapat dilihat pada table 6 pada kolom FP dan FN walaupun sudah di SMOTE masih ada yang salah klasifikasi, bahkan pada kelas minoritas “fear” dan “love” setelah di *oversampling* adanya peningkatan kesalahan jenis False Positif (FP).

Hasil nilai performansi tersebut kemungkinan dipengaruhi oleh dua faktor yaitu jenis model klasifikasi dan kualitas data hasil *preprocessing* yang digunakan. Model klasifikasi dianggap mempunyai pengaruh terhadap nilai performansi karena dari perbandingan hasil pengujian performansi model klasifikasi MNB yang menghasilkan nilai performansi yang lebih baik daripada model klasifikasi *Gaussian Naïve Bayes*. Hal ini terjadi karena model klasifikasi MNB lebih cocok digunakan dalam melakukan klasifikasi teks atau dokumen yang memiliki tipe multi kelas [24]. Kualitas data hasil *preprocessing* juga dianggap mempunyai pengaruh terhadap nilai hasil performansi, dimana kualitas data hasil *preprocessing* pada penelitian ini dianggap masih kurang baik. Metode *convert slangword* pada penelitian ini belum mampu menghasilkan hasil konversi kata *slang* dengan baik karena banyaknya kesalahan ejaan pada dataset yang digunakan sehingga mempengaruhi fitur kata yang dibuat. Jika dilihat dari penelitian M. Adnan Nur dan Nurilmiyanti Wardhani [25] yang memiliki permasalahan yang sama, solusi yang dilakukan untuk mengatasi adanya kesalahan ejaan yaitu dengan menambahkan metode koreksi kesalahan ejaan sebelum melakukan konversi kata *slang*.

Dari hasil pengujian pengaruh nilai k pada SMOTE dari grafik pada gambar 5 dan 6 diketahui bahwa nilai akurasi rata-rata disetiap nilai k selalu sama dan hanya ada peningkatan performansi sebesar 0.01 atau 1% saja pada nilai *f1-score* pada metode gabungan antara MNB dan SMOTE ketika nilai $k = 3$ dan $k = 5$. Sehingga dapat dikatakan bahwa nilai k pada SMOTE tidak mempengaruhi performansi klasifikasi secara signifikan.

4. PENUTUP

Berdasarkan hasil penelitian dapat ditarik kesimpulan bahwa performansi dari gabungan metode *Multinomial Naïve Bayes* (MNB) dan *Synthetic Minority Oversampling Technique* (SMOTE) dalam melakukan analisis emosi pada cuitan Twitter dapat meningkatkan akurasi sebesar 1% dibandingkan tanpa menggunakan SMOTE. Tetapi dalam penelitian ini dapat dilihat bahwa nilai performansi tertinggi dari pengujian performansi yang telah dilakukan didapatkan pada metode gabungan MNB dan SMOTE dengan nilai akurasi rata-rata sebesar 0.65 atau 65% dan nilai *f1-score* rata-rata sebesar 0.66 atau 66%. Hasil nilai performansi tersebut kemungkinan dipengaruhi oleh dua faktor utama yaitu jenis metode klasifikasi dan kualitas data hasil *preprocessing* yang digunakan. Diketahui juga bahwa nilai k pada SMOTE tidak mempengaruhi performansi klasifikasi secara signifikan.

Dari kesimpulan yang telah dijabarkan untuk penelitian selanjutnya disarankan untuk menggunakan model *Multinomial Naïve Bayes* dalam melakukan klasifikasi teks multi kelas agar hasil performansi yang dihasilkan lebih baik, jika menghadapi data yang imbalance dapat menggunakan SMOTE dan juga untuk memperbaiki akurasi dalam penelitian berikutnya dapat menggunakan metode koreksi kesalahan ejaan sebelum melakukan konversi kata *slang*.

DAFTAR PUSTAKA

- [1] S. Dixon, "Leading Countries Based On Number Of Twitter Users As Of January 2022," 2022. [Online]. Available: <https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/>. [Accessed: 09-Nov-2022].
- [2] H. Rachmi, Suparni, And A. Al Kaafi, "Analisis Sentimen Sistem Ganjil Genap Kota Bogor," *J. Tek. Elektro, Teknol. Inf. Dan Komput.*, Vol. 5, No. 2, Pp. 92–99, 2021.
- [3] F. S. Alnaz And W. Maharani, "Analisis Emosi Melalui Media Sosial Twitter Dengan Menggunakan Metode Naïve Bayes Dan Perbandingan Fitur N-Gram Dan Tf-Idf," *eProceedings of Engineering* 8, Pp. 1–14, 2021.
- [4] Rimbun Siringoringo, "Klasifikasi Data Tidak Seimbang Menggunakan Algoritma Smote Dan K-Nearest Neighbor," *J. Isd (Information Syst. Dev.)*, Vol. 3, No. 1, Pp. 44–49, 2018.
- [5] A. Nurlaila, Wiranto, And R. Saptono, "Classification Of Customers Emotion Using Naïve Bayes Classifier (Case Study : Natasha Skin Care)," *Itsart J. Ilm. Teknol. Dan Inf.*, Vol. 6, No. 2, Pp. 92–97, 2017.
- [6] A. A. Arifiyanti And E. D. Wahyuni, "Smote : Metode Penyeimbang Kelas Pada Klasifikasi Data Mining," *J. Teknologi Informasi dan Komukasi*, Vol. Xv, Pp. 34–39, 2020.
- [7] C. Padurariu And M. E. Breaban, "Dealing With Data Imbalance In Text Classification," *Procedia Comput. Sci.*, Vol. 159, Pp. 736–745, 2019.

- [8] S. Chohan, A. Nugroho, A. Maezar, B. Aji, And W. Gata, "Analisis Sentimen Aplikasi Duolingo Menggunakan Metode Naïve Bayes Dan Synthetic Minority Over Sampling Technique," *Paradig. – J. Inform. Dan Komput.*, Vol. 22, No. 2, 2020.
- [9] M. S. Saputri, R. Mahendra, And M. Adriani, "Emotion Classification On Indonesian Twitter Dataset," *Proceeding Int. Conf. Asian Lang. Process*, 2018.
- [10] G. Qiong, W. Xian-Ming, W. Zhao, N. Bing, And C.-S. Xin, "An Improved Smote Algorithm Based On Genetic Algorithm For Imbalanced Data Classification," *J. Digit. Inf. Manag.*, Vol. 14, No. 2, Pp. 92–103, 2016.
- [11] N. Indriani, E. Rainarli, And K. E. Dewi, "Peringkasan Dan Support Vector Machine Pada Klasifikasi Dokumen," *J. Infotel*, Vol. 9, No. 4, Pp. 416–421, 2017.
- [12] A. H. Yunas, Yusra, And M. Fikry, "Klasifikasi Tweet E-Commerce Dengan Menggunakan Metode Support Vector Machine," *J. CoreIT: Jurnal Hasil Penelitian Ilmu Komputer dan Teknologi Informasi*, Vol. 4, No. 2, Pp. 50–55, 2018.
- [13] I. Najiyah And I. Haryanti, "Sentimen Analisis Covid-19 Dengan Metode Probabilistic Neural Network Dan Tf-Idf," *J. Responsif: Riset Sains Dan Informatika*, Vol. 3, No. 1, Pp. 100–111, 2021.
- [14] G. S. Ramadhan, B. Irawan, And C. Setianingsih, "Klasifikasi Emosi Pada Lirik Lagu Menggunakan Algoritma Naïve Bayes Dan Particle Swarm Optimization," *E-Proceeding Eng.*, Vol. 8, No. 5, Pp. 6293–6306, 2021.
- [15] A. Nugroho, "Analisis Sentimen Pada Media Sosial Twitter Menggunakan Naive Bayes Classifier Dengan Ekstrasi Fitur N-Gram," No. 2, Pp. 200–209, 2018.
- [16] D. N. Chandra, G. Indrawan, And I. N. Sukajaya, "Klasifikasi Berita Lokal Radar Malang Menggunakan Metode Naïve Bayes Dengan Fitur N-Gram," *J. Ilmiah Teknologi Informasi Asia*, Vol. 10, No. 1, Pp. 11–19, 2016.
- [17] C. S. K. Aditya And V. R. S. Nastiti, "Sistem Temu Kembali Informasi Buku Hadits Menggunakan Pembobotan Term Frequency Inverse Document Frequency Dan Cosine Similarity," *Sentra(Seminar Nas. Teknol. Dan Rekayasa)*, Vol. 5, No. 2019, Pp. 124–131, 2019.
- [18] N. V Chawla, K. W. Bowyer, L. O. Hall, And W. P. Kegelmeyer, "Smote : Synthetic Minority Over-Sampling Technique," *J. of artificial intelligence research*, Vol. 16, Pp. 321–357, 2002.
- [19] A. M. Ws, I. P. Ananda, M. A. Rizki, Z. D. Hapsari, And R. Nooraeni, "Penerapan Metode Resampling Dalam Mengatasi Imbalanced Data Pada Determinan Kasus Diare Pada Balita Di Indonesia," *J. Matematika dan Statistika serta Aplikasinya*, Vol. 8, No. 1, 2020.
- [20] R. Rinaldi, R. Goejantoro, And Syaripuddin, "Penerapan Metode Klasifikasi Multinomial Naive Bayes (Studi Kasus : Pt Prudential Life Samarinda Tahun 2019)," *Eksponensial*, Vol. 12, No. 2, Pp. 111–118, 2021.
- [21] F. D. A. Nur Ghaniaviyanto Ramadhan, "Sentiment Analysis On Vaccine Covid-19 Using Word Count And Gaussian Naïve Bayes," *Indones. J. Electr. Eng. Comput. Sci.*, Vol. 26, No. 3, Pp. 1765–1772, 2022.
- [22] J. Winahyu And I. Suharjo, "Aplikasi Web Analisis Sentimen Dengan Algoritma Multinomial Naïve Bayes," *KARMAPATI (Kumpulan Artikel Mahasiswa Pendidikan Teknik Informatika)*, Vol. 10, Pp. 206–214, 2021.
- [23] R. R. R. Arisandi, B. Warsito, And A. R. Hakim, "Aplikasi Naïve Bayes Classifier (Nbc) Pada Klasifikasi Status Gizi Balita Stunting Dengan Pengujian K-Fold Cross Validation," *J. Gaussian*, Vol. 11, No. 1, Pp. 130–139, 2022.
- [24] A. S. Andini, D. T. Murdiansyah, And K. M. Lhaksana, "Topic Classification Of Islamic Question And Answer Using Naïve Bayes And Tf-Idf Method," *Comput. Eng. Appl.*, Vol. 10, No. 3, Pp. 151–160, 2021.
- [25] M. A. Nur And N. Wardhani, "Optimasi Normalisasi Kata Pada Data Twitter Untuk Meningkatkan Akurasi Analisis Sentimen (Studi Kasus Respon Masyarakat Terhadap Layanan Teman Bus)," *J. Fokus Elektroda: Energi Listrik, Telekomunikasi, Komputer, Elektronika dan Kendali*, Vol. 07, No. 04, Pp. 237–243, 2022.